CSE 713 - Advanced Syntactic Pattern Recognition
Paper Presentation

Explainable Classifier Supporting Decision-making for Breast Cancer Diagnosis from Histopathological Images

Presented by :
(Group 12)
Mosarrat Rumman - 20266007
Abu Nayeem Tasneem - 20266002
Junaid Bin Kibria - 21166022
Israt Jahan Ritun - 21166038

Introduction

- Highly accurate models lack transparencies in rationalizing their decisions and interpretability to humans.
- Trade-off between transparency and accuracy for classification algorithms
- Algorithm needs to be explainable to humans with high accuracy.

In recent years:

- Lundberg and this co-author introduced in their paper “Explainable machine-learning predictions for the prevention of hypoxaemia during surgery” an algorithm that can predict hypoxaemia and explain contributing factor.
- Sabol and his authors introduced in two of their papers “Cumulative fuzzy class membership criterion decision-based classifier” and “Semantically explainable fuzzy classifier”

Contribution and process overview

- Proposed an improved version of Cumulative Fuzzy Class Membership Criterion (CFCMC).
- Created a semantically explainable algorithm using CFCMC that can classify breast cancer from histopathology samples and interpret additional information about the classification.

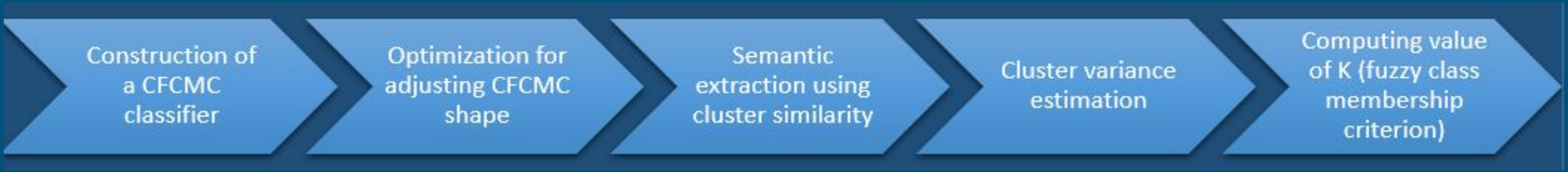


Fig 1: Overview of proposed method

Proposed method

- Cumulative Fuzzy Class Membership Criterion decision based classifier
  - Feature space split into clusters
  - Triangular function defined, eq(1)
  - CFMC is defined as eq(2)

$$\kappa_{\widetilde{p}_{i,j,k}}(\overline{x}) = \begin{cases} 1 - \frac{\|\widetilde{p}_{i,j,k} - \overline{x}\|}{a_{i,j}} & \|\widetilde{p}_{i,j,k} - \overline{x}\| < a_{i,j} \\ 0 & otherwise \end{cases},$$

Eq(1)

$$\chi_{C_i}(\overline{x}) = \max_j \left(\frac{1}{K_{i,j}} \sum_{k=1}^{K_{i,j}} \kappa_{\widetilde{p}_{i,j,k}}(\overline{x}) \right),$$

Eq(2)

$$CL(\overline{x}) = C_{\operatorname{argmax}_i (\chi_{C_i}(\overline{x}))}.$$

Eq(3)

  - Class with maximum CFCMC value is winner.
- Training is divided into 2 parts
  - Initialization: i. Data splitting, ii. Clustering, iii. Parameter initialization
  - Learning is mostly optimization.

Proposed method contd.

- Optimization algorithm such as Repeated annealing, hill climbing, or gradient methods can be used.
- Adjustment of CFCMS's surface shape occurs to achieve highest accuracy.

$$\overline{p}_i = [a_{i,1}, K_{i,1}; \cdots ; a_{i,j}, k_{i,j}; \cdots ; a_{i,n^i_{cl}}, k_{i,n^i_{cl}}] \quad \text{Eq(4)}$$

- Optimization is done to minimize the error rate from validation set.
- Semantics are the relationship between clusters of data.
- Variations of the clusters from the parameters of CFCMC is measured.
- Which is used for finding the maximal membership criterion value of a single cluster.

Proposed method contd.

- ➢ Computation of K $k = p_1 \frac{K_{Cl}}{m_{Cl}} + p_2$ Eq(5)
- ➢ Euclidean distance displaced by custom distance

$$d = \frac{1}{n} \sum_{i=1}^{n} | x_{A,i} - x_{B,i} |$$ Eq(6)

- ➢ Custom Distance achieved stability with K parameter

Experiment

- ➢ Performance comparison with CNN, SAE and Deep multi-layered perceptron
- ➢ Dataset: 277,524 50x50 pixel RGB digital image patches
- ➢ Experimental setup of CFCMC, CNN, SAE and Deep multi-layered perceptron
- ➢ Performance results :

COMPARISON OF CFCMC WITH THREE CONTENDING CLASSIFIERS

CFCMC	MLP	CNN	SAE
78.65% ± 1.81%	82.54% ± 1.01%	83.29% ± 0.99%	82.29 % ± 1.97%

Proposed Explanation

- Four Possible Situations for Binary Classifier:
  - Classified as IDC with low possibility for misclassification
  - Classified as IDC with high possibility for misclassification
  - Classified as non-IDC with low possibility for misclassification
  - Classified as non-IDC with high possibility for misclassification

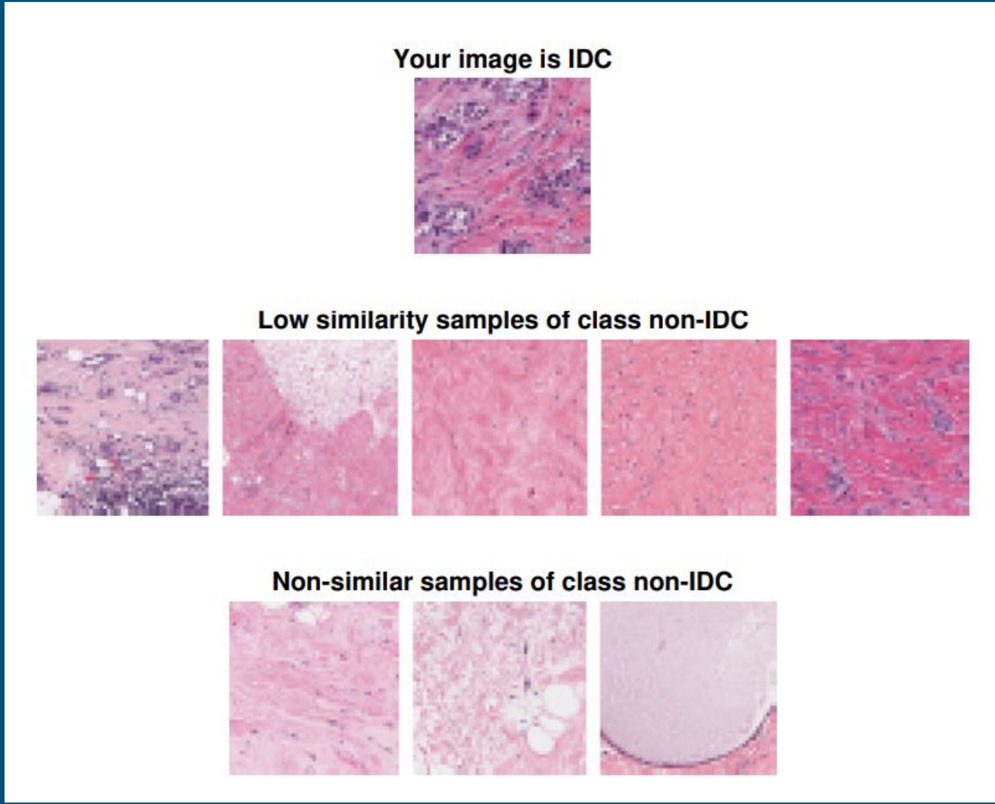
TABLE II
VALUES OF SIMILARITY OF IDC CLUSTERS TO NON-IDC CLUSTERS

Cluster name	non-IDC_1	non-IDC_2	non-IDC_3	non-IDC_4	non-IDC_5	non-IDC_6	non-IDC_7	non-IDC_8
IDC_1	no	no	no	no	no	no	no	low
IDC_2	low	no	low	low	no	low	no	low
IDC_3	low	medium	medium	high	medium	low	low	no
IDC_4	low	no	no	no	no	medium	no	low
IDC_5	low	no	low	low	no	no	no	no
IDC_6	low	high	low	low	medium	no	low	no
IDC_7	medium	low	low	medium	low	medium	no	no
IDC_8	low	no	no	no	no	low	no	no

Proposed Explanation contd.

- Example of Provided Semantics:

  "Your sample is IDC with high degree of confidence, because it has low similarity or non-similar with the non-IDC samples."



THANK YOU!!!!